

(19) European Patent Office

(11) Publication number: 0 276 723 B1

(12)

## EUROPEAN PATENT

(46) Publication date of the patent: 8 Dec. 93

(51) Int. Cl.<sup>5</sup>: C07K 13/00, C12Q 1/68,  
G01N 33/53

(21) Application number: 88100647.2

(22) Application date: 19 Jan. 88

---

(54) Precursor protein of the APC polypeptide, DNA coding for it, and diagnostic use of the DNA and the protein.

---

(30) Priority: 30 Jan. 87 DE 3702/89

(73) Patent holder: BAYER AG

(43) Publication date of the application:  
3 Aug. 88 Patent Journal 88/31

D-51368 Leverkusen (DE)

(45) Notification of the patent issuance:  
8 Dec. 93 Patent Journal 93/49(72) Inventor: Müller-Hill, Benno, Prof. Dr.  
Händelstrasse 53

D-5000 Cologne (DE)

(84) Designated signatory countries:  
AT BE CH DE ES FR GB GR IT LI NL SEInventor: Kang, Jie, Dipl.-Biol.  
Rheinallee 37

D-5300 Bonn 2 (DE)

(56) Citations:  
EP-A-0 274 826  
US-A-4 866 829Inventor: Lemaire, Hans-Georg, Dipl.-Biol.  
Meister Gerhard Strasse 33

D-5000 Cologne (DE)

Inventor: Unterbeck, Axel, Dr.

1 Campbell Avenue,

Apt. 116

West Haven, CT 06516 (US)

NATURE Volume 325, 19 Feb. 1987; J.  
KANG et al., pages 733-736&NUM;SCIENCE Volume 235, 20 Feb. 1987; D.  
GOLDGABER et al., pages 880-887&NUM;PROCEEDINGS OF THE NATL. ACADEMY  
OF SCIENCES USA, Volume 84, June 1987;  
N.K. ROBAKIS et al.; pages 4190-4194&NUM;

---

Note: Within nine months of notification of the issuance of the European Patent anyone may file an opposition to the issued European Patent at the European Patent Office. The opposition shall be submitted in writing, giving reasons. It counts as filed when the opposition fee has been paid (Art. 99(1) European Patent Convention).

---

PROCEEDINGS OF THE NATL. ACADEMY  
OF SCIENCES USA, Volume 83, April 1986;  
A. ROHER et al.; pages 2662-2666&NUM;

PROCEEDINGS OF THE NATL. ACADEMY  
OF SCIENCES USA, Volume 82, June 1985;  
C.L. MASTERS et al.; pages 4245-4249&NUM;

CHEMICAL ABSTRACTS, Volume 104, 1986,  
Columbus, OH (US); C.L. MASTERS et al.,  
page 506, No. 127641f&NUM;

---

## Specification

The present invention concerns the precursor protein of the "amyloid plaque core" (APC) polypeptide, fragments of the precursor protein, and the diagnostic use of the precursor protein or the fragments. Furthermore, the invention concerns the DNA coding for the precursor protein, fragments of this DNA, and the diagnostic use of the DNA and the fragments.

Alzheimer's disease was described for the first time as an independent clinical and pathological phenomenon in 1907 by the German neurologist Alois Alzheimer (Alzheimer, A. (1907) *Zentralblatt für Nervenheilkunde und Psychiatrie*, 177-179). It is the most frequent degenerative brain disease in elderly people. In America alone, approximately 2 million people today suffer from the disease and at least 100,000 of these die each year (Wurtman, R.J. (1985) *Sci. Am.* 252, 48-56).

The disease occurs in people between the age of 40 and 80. Those affected gradually lose their memory and their ability to concentrate. The state of mental decay increases over the course of 3 to 10 years, until the patients neither speak or think, nor can they take care of themselves, and they ultimately die. The cause of this dementia is unknown. There is neither a definitive diagnosis nor a therapy.

Typical changes can be recognized under the microscope in brain autopsies of those who have died from Alzheimer's disease.

- The number of neurons has diminished, especially in the parietal lobes, i.e., in the portions of the brain where the faculties of memory are localized. A loss of neurons which normally release acetylcholine is also clearly noticeable.
- Furthermore, three extremely unusual structures occur in the cerebral cortex, which do not exist in the brain of healthy people and therefore are used for diagnosis (even postmortem):

1) intracellular neurofibrils (NFTs, "neurofibrillary tangles")

In the cell body of neurons of the cerebral cortex and the hippocampus one finds bundles consisting of two filaments helically twisted about each other (PHFs, "paired helical filaments").

2) extracellular amyloid plaques (APC, "amyloid plaque core")

The neuritic plaques contain amyloid and the residue of dead cells, which are scattered in the cerebral cortex, the hippocampus, and the amygdaloid nucleus.

The number of plaques correlates with the degree of mental deterioration.

3) cerebrovascular amyloid (ACA, "amyloid congophilic angiopathy")

By amyloid is meant a protein-rich mass. Such formless protein aggregates are found around the blood vessels or in the wall of blood vessels in the brain.

The protein component of the ACA has been isolated and sequenced (Glennner, G.G. & Wong, C.W. (1984) *Biochem. Biophys. Res. Commun.* 120, 885-890). The amino acid sequence has no homology with known protein sequences. The protein components of the PHF and APC have likewise been isolated and sequenced (Masters, C.L., Multhaupt, G., Simms, G., Potgiesser, J., Martins R.N. and Beyreuther, K. (1985) *EMBO* 4, 2757-2763 and Masters, C.L., Simms, G., Weinman, N.A., Multhaupt, G., McDonald, B.L. and Beyreuther, K. (1985) *Proc. Natl. Acad. Sci. USA* 82, 4245-4249). Based on the amino acid sequences, the same polypeptide is probably involved in all three instances with a

molecular weight of 4.5 kD. The corresponding sequence is shown in figure 1 a-c (position 597-638).

There are several hypotheses about the origin of this APC protein. It might be a normal protein in the brain (or also in another organ), with either runaway biosynthesis regulation or disturbed physiological decomposition. Its build-up in very large quantity might then be the cause of the disease. If it is an abnormal protein and its unusual aggregation ability causes the disease, then it could also be coded by a healthy human gene that has been wrongly controlled by some factor, such as virus, nutrients, or environmental toxins. The error might also involve a modification of the original protein precursor. On the other hand, a viral gene might also be responsible for the synthesis of the APC protein.

It was now attempted to clarify the origin and nature of the APC protein whose aggregation in the cerebral cortex is one of the main biochemical symptoms of Alzheimer patients, in order to thus obtain a tool for improved diagnosis of Alzheimer's disease.

For this, a fetal human brain c-DNA bank with pA + mRNA of the cerebral cortex was constructed.

The c-DNA synthesis was carried out according to Okayama and Berg (Okayama, H. and Berg, P. Mol. Cell. Biol. 2, 161-170 (1982); Okayama, H. and Berg, P. Mol. Cell. Biol. 3, 280-289 (1983)) and the c-DNA was transformed in *E. coli* HB 101 (Aviv, H. and Leder, P. Proc. Natl. Acad. Sci. USA 69, 1408 (1972)). Each of the thus obtained c-DNA banks contains more than  $1 \times 10^6$  independent c-DNA clones.

For sorting through the bank, a DNA probe was used, whose sequence was derived from the sequence of the APC polypeptide. The sequence which corresponds to the amino acids in position 10-16 of the APC was chosen. The corresponding sequence is marked by brackets in figure 1 c (position 1815-1835). To ensure an optimal hybridization, the degeneration of the genetic code was taken into account and a mixture with the following sequence was prepared and used as probe:

5' TT<sup>T</sup>TG<sup>A</sup>ATG<sup>A</sup>ATGCAC<sup>T</sup>TC<sup>A</sup>TA3  
C G G C GG

this is a 64-fold degenerated 20-mer. The corresponding sequence is marked by brackets in figure 1 c (position 1815-1835). A test of 100,000 c-DNA clones of the human fetal cerebral cortex bank led to the isolation of a complete (full-length) c-DNA clone with serial number EC 9.110, which codes for a protein that contains the APC sequence and thus represents the precursor protein of the APC peptide. The sequence of the c-DNA and the amino acid sequence of the coded protein are shown in figure (1). The sequence analysis was done by the Dideoxy Method (Sanger, F., Nicklen, S. and Coulson, A.R. Proc. Natl. Acad. Sci. USA 74, 5463-5467 (1977) and Guidelines for quick and simple plasmid sequencing, Handbook, (1986) Boehringer Mannheim GmbH, Biochemica., D-6800 Mannheim). As of yet, nothing is known about the natural function of the APC precursor protein.

Thus, the present invention concerns the deoxyribonucleic acid of sequence per figure 1 and its functional equivalents. The term functional equivalents in this connection signifies that individual nucleotides within the sequence can be exchanged or derivatized,

due to the degeneration of the genetic code, without this having an influence on the function of the nucleic acid. In particular, the invention concerns the DNA of the sequence per figure 1 from position 1 to position 2089 and its functional equivalents. This part of the DNA is the part which codes for the precursor protein. Due to certain features in the sequence, the protein and the corresponding DNA sequence is an interesting tool for the diagnosis of Alzheimer's disease on a molecular level. Here, in particular, the region from around position 600 to around position 900 is noteworthy. This part codes for an unusually large number of acidic amino acids, in terms of the length of this segment. Also quite conspicuous are the seven threonines (position: DNA 819-840/amino acids 274-280) in succession. Such regions are especially interesting for the development of DNA probes for diagnostics, since they are unique in their unusual sequence and thus allow a highly specific detection. The invention also concerns fragments of DNA from figure 1 or oligonucleotides derived from this DNA and their use as probes in diagnostics. For hybridization experiments, the DNA will not be used in its full length. Normally, for hybridization one uses fragments with a length of around 10 to 50 nucleotides. Longer fragments usually lead to problems of manipulation. Fragments with fewer than 10 nucleotides usually do not have sufficient specificity or the bonding is too weak.

The DNA per figure 1 or the fragments of this DNA can be used very well during diagnosis of Alzheimer's disease, for determining mutations such as deletions, insertions and point mutations or rearrangement mistakes.

The present invention enables a diagnosis of Alzheimer's disease on the molecular level. This also holds for presymptomatic diagnosis of Alzheimer's disease. The analysis can be performed with known techniques of DNA technology, such as the techniques described by Antonarkals et al. (1985), in Hum. Gen. 69, 1-14.

The present invention also includes the precursor protein coded by the DNA or its fragments. The detection of this protein or its fragments likewise represents a starting point for the diagnosis of Alzheimer's disease. Here, again, the peculiarities of the sequence (amino acids: roughly position 200 to position 290) are of special significance. Fragments of the precursor protein, especially from the region of 200 to 290, can be used very well as antigen peptides for the preparation of polyclonal or monoclonal antibodies, which in turn can be used in diagnostics.

Functional equivalents in connection with the protein or the peptides means that variations in the form of amino acid exchanges or derivatization are possible both in the sequence of the protein and in the peptides, which have no influence on the function of this peptide as an antigen, for example.

Legend for Figure 1 a-c

Nucleotide sequence 5'  $\rightarrow$  3' of the c-DNA clone which codes for the precursor protein of the APC polypeptide and the amino acid sequence derived from the DNA. The amino acids are designated by the following single-letter codes:

## Amino acids

A Ala alanine  
B Asz Asn or Asp  
C Cys cysteine (cystine)  
D Asp asparaginic acid  
E Glu glutaminic acid  
F Phe phenylalanine  
G Gly glycerin  
H His histidine  
    HS homoserine  
    HSL homoserine lactone  
I Ile isoleucine  
K Lys lysine  
L Leu leucine  
M Met methionine  
N Asp asparagine  
    Nle norleucine  
P Pro proline  
Q Gln glutamine  
R Arg arginine  
S Ser serine  
T Thr threonine  
V Val valine  
W Trp tryptophan  
Y Tyr tyrosine  
Z Glx Glu or Gln  
X not identified

19



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



11 Veröffentlichungsnummer: **0 276 723 B1**

12

## EUROPÄISCHE PATENTSCHRIFT

13 Veröffentlichungstag der Patentschrift: 08.12.93

11 Int. Cl.<sup>5</sup>: **C07K 13/00, C12Q 1/68,  
G01N 33/53**

21 Anmeldenummer: 88100647.2

22 Anmeldetag: 19.01.88

54 Vorläuferprotein des APC-Polypeptids, dafür codierende DNA und diagnostische Verwendung der DNA und des Proteins.

30 Priorität: 30.01.87 DE 3702789

43 Veröffentlichungstag der Anmeldung:  
03.08.88 Patentblatt 88/31

45 Bekanntmachung des Hinweises auf die  
Patenterteilung:  
08.12.93 Patentblatt 93/49

64 Benannte Vertragsstaaten:  
AT BE CH DE ES FR GB GR IT LI NL SE

56 Entgegenhaltungen:  
EP-A- 0 274 826  
US-A- 4 666 829

NATURE, Band 325, 19. Februar 1987; J.  
KANG et al., Seiten 733-736&NUM;

SCIENCE, Band 235, 20. Februar 1987; D.  
GOLDGABER et al., Seiten 880-887&NUM;

PROCEEDINGS OF THE NATL. ACADEMY OF  
SCIENCES USA, Band 84, Juni 1987; N.K.  
ROBAKIS et al.; Seiten 4190-4194&NUM;

73 Patentinhaber: BAYER AG

D-51368 Leverkusen(DE)

72 Erfinder: MÜLLER-HILL, Benno, Prof. Dr.  
Händelstrasse 53

D-5000 Köln(DE)

Erfinder: Kang, Jie, Dipl.-Biol.

Rheinallee 37

D-5300 Bonn 2(DE)

Erfinder: Lemaire, Hans-Georg, Dipl.-Biol.

Meister Gerhard Strasse 33

D-5000 Köln 1(DE)

Erfinder: Unterbeck, Axel, Dr.

1 Campbell Avenue,

Apt. 116

West Haven, CT 06516(US)

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99(1) Europäisches Patentübereinkommen).

EP 0 276 723 B1

EP 0 276 723 B1

PROCEEDINGS OF THE NATL. ACADEMY OF  
SCIENCES USA, Band 83, April 1986; A. RO-  
HER et al., Seiten 2662-2666&NUM;

PROCEEDINGS OF THE NATL. ACADEMY OF  
SCIENCES USA, Band 82, Juni 1985; C.L. MA-  
STERS et al.; Seiten 4245-4249&NUM;

CHEMICAL ABSTRACTS, Band 104, 1986, Co-  
lumbus, OH (US); C.L. MASTERS et al., Seite  
506, Nr. 127641f&NUM;



## Beschreibung

Die vorliegende Erfindung betrifft das Vorläuferprotein des "amyloid plaque core" (APC)-Polypeptids, Fragmente des Vorläuferproteins sowie die diagnostische Verwendung des Vorläuferproteins bzw. der Fragmente. Weiterhin betrifft die Erfindung die für das Vorläuferprotein codierende DNA, Fragmente dieser DNA sowie die diagnostische Verwendung der DNA bzw. der Fragmente.

Die Alzheimersche Krankheit wurde im Jahre 1907 von dem deutschen Neurologen Alois Alzheimer zum ersten Mal als eine eigenständige klinische und pathologische Erscheinung beschrieben (Alzheimer, A. (1907) Zentralblatt für Nervenheilkunde und Psychiatrie, 177-179). Sie ist die häufigste degenerative Gehirnerkrankung bei alten Menschen. Allein in Amerika leiden ungefähr 2 Millionen Menschen heute an der Krankheit und mindestens 100 000 davon sterben jedes Jahr (Wurtman, R.J. (1985) Sci.Am. 252, 48-56).

Die Krankheit tritt beim Menschen zwischen dem 40. und 80. Lebensjahr auf. Die Betroffenen verlieren allmählich ihr Gedächtnis und ihre Konzentrationsfähigkeit. Der Zustand des geistigen Verfalls steigt sich innerhalb von 3 bis 10 Jahren, bis die Patienten weder sprechen, denken noch für sich sorgen können und schließlich sterben. Die Ursache für diese Demenz ist unbekannt. Es gibt weder eine definitive Diagnose noch eine Therapie.

Bei Hirnautopsien von an der Alzheimerschen Krankheit gestorbenen Menschen sind im Mikroskop typische Veränderungen zu erkennen.

- Die Zahl der Neuronen hat abgenommen, vor allem in den Scheitellappen, also in den Teilen des Gehirns, wo die Gedächtnisleistungen lokalisiert sind. Ein Verlust an Neuronen, die normalerweise Acetylcholin freisetzen, ist ebenfalls deutlich sichtbar.

- Weiterhin kommen im cerebralen Cortex drei äußerst ungewöhnliche Strukturen vor, die im Gehirn gesunder Menschen nicht existieren und daher zur Diagnose (nach dem Tod) benutzt werden:

1) intrazelluläre Neurofibrillen (NFTs, "neurofibrillary tangles")

Im Zellkörper von Neuronen der Großhirnrinde und des Hippocampus findet man Bündel, die aus zwei helixartig umeinandergedrehten Filamenten (PHFs, "paired helical filaments") bestehen.

2) extrazelluläre amyloide Plaques (APC, "amyloid plaque core")

Die neuritischen Plaques enthalten Amyloid und die Überreste abgestorbener Zellen, sie sind in der Großhirnrinde, dem Hippocampus und dem Mandelkern verstreut. Die Anzahl der Plaques korreliert mit dem Grad des gei-

stigen Verfalls.

3) cerebrovasculäres Amyloid (ACA, "amyloid congophilic angiopathy")

Als Amyloid bezeichnet man eine proteinreiche Masse. Solche formlosen Proteinaggregate sind ringsum die Blutgefäße bzw. in der Wand von Blutgefäßen im Gehirn zu finden.

Die Proteinkomponente des ACA wurde isoliert und sequenziert (Glenner, G.G. & Wong, C.W. (1984) Biochem. Biophys. Res. Commun. 120, 885-890). Die Aminosäuresequenz hat keine Homologie zu bekannten Proteinsequenzen. Die Proteinkomponenten der PHFs bzw. APC wurden ebenfalls isoliert und sequenziert (Masters, C.L., Multhaup, G., Simms, G., Pottgiesser, J., Martins R.N. and Beyreuther, K. (1985) EMBO 4, 2757-2763 und Masters, C.L., Simms, G., Weinman, N.A., Multhaup, G., McDonald, B.L. and Beyreuther, K. (1985) Proc. Natl. Acad. Sci. USA 82, 4245-4249). Nach den Aminosäuresequenzen handelt es sich wahrscheinlich in allen drei Fällen um das gleiche Polypeptid mit einem Molekulargewicht von 4,5 kD. Die entsprechende Sequenz ist in der Fig. 1 a-c eingerahmt (Position 597-638).

Es gibt mehrere Hypothesen um die Herkunft dieses APC-Proteins zu erklären. Es könnte ein normales Protein im Gehirn (oder auch in einem anderen Organ) sein, bei dem entweder die Regulation der Biosynthese entgleist oder der physiologische Abbau gestört ist. Die Anhäufungen in sehr großer Menge könnten dann die Ursache der Krankheit sein. Falls es ein abnormes Protein ist und seine ungewöhnliche Aggregationsfähigkeit die Krankheit verursacht, könnte es ebenfalls von einem gesunden, menschlichem Gen kodiert werden, das durch irgendeinen Faktor, z.B. Viren, Nahrungsmittel oder Umweltgifte falsch gesteuert würde. Der Fehler könnte auch in einer Modifikation der ursprünglichen Protein-Vorstufe bestehen. Andererseits könnte aber auch ein virales Gen für die Synthese des APC-Proteins verantwortlich sein.

Es wurde nun versucht die Herkunft und Natur des APC-Proteins zu klären, dessen Aggregation im cerebralen Cortex eines der biochemischen Hauptsymptome von Alzheimer-Patienten ist, um damit ein Werkzeug zur verbesserten Diagnostik der Alzheimer-Krankheit zu erhalten.

Dazu wurde eine fötale, humane Hirn c-DNA-Bank mit pA+mRNA des cerebralen Cortex konstruiert.

Die c-DNA-Synthese wurde nach Okayama und Berg (Okayama, H. and Berg, P. Mol. Cell. Biol. 2, 161-170 (1982); Okayama, H. and Berg, P. Mol. Cell. Biol. 3, 280-289 (1983)) durchgeführt und die c-DNA in E. coli HB 101 transformiert (Aviv, H. and Leder, P. Proc. Natl. Acad. Sci. USA 69, 1408 (1972)). Jede der so erhaltenen c-DNA-Banken enthält mehr als  $1 \times 10^6$  unabhängige c-DNA-Klone.

Für das Sichten der Bank wurde eine DNA-Sonde benutzt, deren Sequenz aus der Sequenz des APC-Polypeptids abgeleitet war. Es wurde die Sequenz gewählt, die den Aminosäuren in Position 10 - 16 des APC entspricht. Die entsprechende Sequenz ist in der Fig. 1 c durch eine Klammer markiert (Position 1815-1835). Um eine optimale Hybridisierung zu gewährleisten, wurde die Degeneration des genetischen Codes berücksichtigt und als Sonde ein Gemisch mit der folgenden Sequenz

5' TT<sup>T</sup> TG<sup>A</sup> TG<sup>A</sup> TGCAC<sup>T</sup> T<sup>CA</sup> TA3' ,  
C G G C GG

hergestellt und benutzt. Dies ist ein 64fach degeneriertes 20-mer. Die entsprechende Sequenz ist in der Fig. 1 c durch eine Klammer markiert (Position 1815-1835). Ein Test von 100 000 c-DNA Klonen der menschlichen fötalen cerebralen Cortex Bank führte zu der Isolierung eines vollständigen (full-length) c-DNA-Klons mit der Serien-Nr. EC 9.110, der ein Protein kodiert, welches die APC-Sequenz enthält und somit das Vorläufer-Protein des APC-Peptids darstellt. Die Sequenz der c-DNA und die Aminosäuresequenz des kodierten Proteins sind der Figur (1) zu entnehmen. Die Sequenzanalyse erfolgte nach der Dideoxy-Methode (Sanger, F., Nicklen, S. and Coulson, A.R. Proc. Natl. Acad. Sci. USA 74, 5463-5467 (1977) und Guidelines for quick and simple Plasmid Sequencing, Handbook, (1986) Boehringer Mannheim GmbH, Biochemica, D-6800 Mannheim). Über die natürliche Funktion des APC-Vorläuferproteins ist zur Zeit noch nichts bekannt.

Die vorliegende Erfindung betrifft also die Desoxyribonucleinsäure der Sequenz gemäß Fig. 1 und deren funktionellen Äquivalente. Der Ausdruck funktionelle Äquivalente bedeutet in diesem Zusammenhang, daß innerhalb der Sequenz, bedingt durch die Degeneration des genetischen Code, einzelne Nucleotide ausgetauscht oder auch derivatisiert werden können, ohne daß dies einen Einfluß auf die Funktion der Nukleinsäure hat. Insbesondere betrifft die Erfindung die DNA der Sequenz gemäß Fig. 1 von Position 1 bis Position 2089 und deren funktionellen Äquivalente. Dieser Teil der DNA ist der Teil, der für das Vorläuferprotein codiert. Durch einige Besonderheiten in der Sequenz ist das Protein und die korrespondierende DNA-Sequenz ein interessantes Werkzeug für die Diagnose der Alzheimer Krankheit auf molekularer Ebene. Hier ist insbesondere der Bereich von etwa Position 600 bis etwa Position 900 erwähnenswert. Dieser Teil codiert für eine ungewöhnlich große Zahl von sauren Aminosäuren, bezogen auf die Länge dieses Abschnitts. Ganz besonders auffällig sind auch die sieben Threonine (Position: DNA 819

-840 / Aminosäuren 274 - 280) hintereinander. Für die Entwicklung von DNA-Sonden für die Diagnostik sind solche Bereiche besonders interessant, da sie durch ihre außergewöhnliche Sequenz einzigartig sind und somit eine hochspezifische Detektion erlauben. Die Erfindung betrifft auch Fragmente der DNA aus Fig. 1 bzw. aus dieser DNA abgeleitete Oligonukleotide sowie deren Verwendung als Sonden in der Diagnostik. Für Hybridisierungsexperimente wird die DNA nicht in der vollen Länge eingesetzt. Normalerweise benutzt man für Hybridisierungen Fragmente einer Länge von etwa 10 bis 50 Nukleotiden. Längere Fragmente führen meist zu Problemen bei der Handhabung. Fragmente mit weniger als 10 Nukleotiden besitzen meistens nicht die hinreichende Spezifität bzw. die Bindung ist zu schwach.

Die DNA gemäß Fig. 1 bzw. die Fragmente dieser DNA können sehr gut bei der Diagnose der Alzheimer-Erkrankung verwendet werden, zum Feststellen von Mutationen wie z.B. von Deletionen, Insertionen und Punktmutationen oder Rearrangement-Fehlern.

Die vorliegende Erfindung ermöglicht die Diagnostik der Alzheimer-Erkrankung auf molekularer Ebene. Dies gilt gleichermaßen für die Präsymptomatische Diagnostik der Alzheimer-Erkrankungen. Durchgeführt werden kann die Analytik mit bekannten Techniken aus der DNA-Technologie wie z.B. den von Antonarkais et al (1985), in Hum. Gen 69, 1-14 beschriebenen Techniken.

Zu der vorliegenden Erfindung gehört auch das von der DNA codierte Vorläuferprotein bzw. dessen Fragmente. Der Nachweis dieses Proteins bzw. der Fragmente stellt ebenfalls einen Ansatzpunkt für die Diagnose der Alzheimer Krankheit dar. Auch hier sind die Besonderheiten der Sequenz (Aminosäuren: ca. Position 200 bis ca. Position 290) von besonderer Bedeutung. Fragmente des Vorläuferproteins, insbesondere aus der Region 200 bis 290, können sehr gut als antigene Peptide für die Herstellung von polyklonalen oder monoklonalen Antikörpern verwendet werden, die wiederum Verwendung in der Diagnostik finden.

Funktionelle Äquivalente im Zusammenhang mit dem Protein bzw. den Peptiden bedeutet, daß sowohl in der Sequenz des Proteins als auch bei den Peptiden Variationen in Form von Aminosäureaustauschen bzw. Derivatisierungen möglich sind, die keinen Einfluß auf die Funktion dieser Peptide z.B. als Antigen haben.

Legende zu Figur 1 a-c.

Nukleotid-Sequenz 5' → 3' des c-DNA-Klons, der für das Vorläuferprotein des APC-Polypeptids codiert und die aus der DNA abgeleitete Aminosäuresequenz. Die Aminosäuren sind mit dem fol-

EP 0 276 723 B1

6

genden Einbuchstaben-Code bezeichnet:

#### Aminosäuren

A	Ala Alanin
B	Asz Asn oder Asp
C	Cys Cystein (Cystin)
D	Asp Asparaginsäure
E	Glu Glutaminsäure
F	Phe Phenylalanin
G	Gly Glycerin
H	His Histidin
	HS Homoserin
	HSL Homoserinlacton
I	Ile Isoleucin
K	Lys Lysin
L	Leu Leucin
M	Met Methionin
N	Asn Asparagin
	Nle Norleucin
P	Pro Prolin
Q	Gln Glutamin
R	Arg Arginin
S	Ser Serin
T	Thr Threonin
V	Val Valin
W	Trp Tryptophan
Y	Tyr Tyrosin
Z	Glx Glu oder Gln
X	nicht identifiziert

#### Patentansprüche

Patentansprüche für folgende Vertragsstaaten

: AT, BE, CH, DE, FR, GB, IT, LI, NL, SE

1. Desoxyribonukleinsäure der Sequenz gemäß Fig. 1a - 1c und solche funktionellen Equivalente dieser Sequenz, in denen, bedingt durch die Degeneration des genetischen Codes, einzelne Nukleotide ausgetauscht oder derivatisiert sind, ohne daß dies einen Einfluß auf die Funktion der Nukleinsäure hat.

2. Desoxyribonukleinsäure der Sequenz von Position 1 bis Position 2089 gemäß Fig. 1a - 1c und solche funktionellen Equivalente dieser Sequenz, in denen, bedingt durch die Degeneration des genetischen Codes, einzelne Nukleotide ausgetauscht oder derivatisiert sind, ohne daß dies einen Einfluß auf die Funktion der Nukleinsäure hat.

3. Desoxyribonukleinsäure der Sequenz von Position 600 bis 900 aus Fig. 1a - 1c und solche funktionellen Equivalente dieser Sequenz, in denen, bedingt durch die Degeneration des genetischen Codes, einzelne Nukleotide ausgetauscht oder derivatisiert sind, ohne daß dies

einen Einfluß auf die Funktion der Nukleinsäure hat.

4. Verwendung der Desoxyribonukleinsäure gemäß den Ansprüchen 1 bis 3 zur in vitro Diagnose der Alzheimer-Erkrankung.

5. Testkit enthaltend die Desoxyribonukleinsäure gemäß den Ansprüchen 1 bis 3.

6. Vorläuferprotein des Amyloiden Plaque Proteins der Sequenz gemäß Fig. 1a - 1c und solche funktionellen Equivalente dieser Sequenz, in denen, bedingt durch die Degeneration des genetischen Codes, einzelne Aminosäuren ausgetauscht oder derivatisiert sind, ohne daß dies einen Einfluß auf die Funktion des Proteins hat.

7. Protein der Sequenz von Position 210 bis Position 290 gemäß Fig. 1a - 1c und solche funktionellen Equivalente, in deren Sequenz, bedingt durch die Degeneration des genetischen Codes, einzelne Aminosäuren ausgetauscht oder derivatisiert sind, ohne daß dies einen Einfluß auf die Funktion des Proteins hat.

8. Fragmente der Proteine nach den Ansprüchen 6 und 7.

9. Verwendung der Proteine bzw. Fragmente gemäß den Ansprüchen 6 bis 8 zur in vitro Diagnose der Alzheimer Erkrankung.

10. Nicht therapeutische Verwendung der Proteine bzw. Fragmente gemäß den Ansprüchen 6 bis 8 als Antigene zur Herstellung von Antikörpern.

11. Antikörper gerichtet gegen die Proteine bzw. Fragmente gemäß den Ansprüchen 6 bis 8.

12. Testkit enthaltend die Antikörper gemäß Anspruch 11.

13. Verwendung der Antikörper gemäß Anspruch 11 zur in vitro Diagnose der Alzheimer Erkrankung.

Patentansprüche für folgende Vertragsstaaten : ES, GR

1. Verwendung der Desoxyribonukleinsäure gemäß Fig. 1a - 1c und solche funktionellen Equivalente dieser Sequenz, in denen, bedingt durch die Degeneration des genetischen Codes, einzelne Nukleotide ausgetauscht oder derivatisiert sind, ohne daß dies einen Einfluß auf die Funktion der Nukleinsäure hat, zur in

in vitro Diagnose der Alzheimer-Erkrankung.

2. Verwendung der Desoxyribonukleinsäure der Sequenz von Position 1 bis Position 2089 gemäß Fig. 1a - 1c und solche funktionellen Äquivalente dieser Sequenz, in denen, bedingt durch die Degeneration des genetischen Codes, einzelne Nukleotide ausgetauscht oder derivatisiert sind, ohne daß dies einen Einfluß auf die Funktion der Nukleinsäure hat, zur in vitro Diagnose der Alzheimer-Erkrankung.
3. Verwendung der Desoxyribonukleinsäure der Sequenz von Position 600 bis Position 900 aus Fig. 1a - 1c und solche funktionellen Äquivalente dieser Sequenz, in denen, bedingt durch die Degeneration des genetischen Codes, einzelne Nukleotide ausgetauscht oder derivatisiert sind, ohne daß dies einen Einfluß auf die Funktion der Nukleinsäure hat, zur in vitro Diagnose der Alzheimer-Erkrankung.
4. Verwendung des Vorläuferproteins des Amyloiden Plaque Proteins der Sequenz gemäß Fig. 1a - 1c und solche funktionellen Äquivalente dieser Sequenz, in denen, bedingt durch die Degeneration des genetischen Codes, einzelne Aminosäuren ausgetauscht oder derivatisiert sind, ohne daß dies einen Einfluß auf die Funktion des Proteins hat, zur in vitro Diagnose der Alzheimer-Erkrankung.
5. Verwendung der Proteine bzw. Fragmente gemäß den Ansprüchen 1 bis 4 als Antigene zur Herstellung von Antikörpern.
6. Verwendung der Antikörper gemäß Anspruch 5 zur in vitro Diagnose der Alzheimer-Erkrankung.

#### Claims

Claims for the following Contracting States :  
AT, BE, CH, DE, FR, GB, IT, LI, NL, SE

1. Deoxyribonucleic acid of the sequence shown in Figures 1a - 1c and those functional equivalents of the sequence in which, due to the degeneration of the genetic code, individual nucleotides are exchanged or derivatised without this having an effect on the function of the nucleic acid.
2. Deoxyribonucleic acid of the sequence from position 1 to position 2089 as shown in Figures 1a - 1c and those functional equivalents of this sequence in which, due to the degeneration of the genetic code, individual nucleotides are exchanged or derivatised without this having

an effect on the function of the nucleic acid.

3. Deoxyribonucleic acid of the sequence from position 600 to position 900 as shown in Figures 1a - 1c and those functional equivalents of this sequence in which, due to the degeneration of the genetic code, individual nucleotides are exchanged or derivatised without this having an effect on the function of the nucleic acid.
4. Use of the deoxyribonucleic acid according to Claims 1 to 3 for the in vitro diagnosis of Alzheimer's disease.
5. Test kit containing the deoxyribonucleic acid according to Claims 1 to 3.
6. Precursor protein of amyloid plaque protein of the sequence shown in Figures 1a - 1c and those functional equivalents of this sequence in which, due to the degeneration of the genetic code, individual amino acids are exchanged or derivatised without this having an effect on the function of the protein.
7. Protein of the sequence from position 210 to position 290 as shown in Figures 1a - 1c and those functional equivalents in whose sequence, due to the degeneration of the genetic code, individual amino acids are exchanged or derivatised without this having an effect on the function of the protein.
8. Fragments of the proteins according to Claims 6 and 7.
9. Use of the proteins or fragments according to Claims 6 to 8 for the in vitro diagnosis of Alzheimer's disease.
10. Non-therapeutic use of the proteins or fragments according to Claims 6 to 8 as antigens for the production of antibodies.
11. Antibodies directed against the proteins or fragments according to Claims 6 to 8.
12. Test kit containing the antibodies according to Claim 11.
13. Use of the antibodies according to Claim 11 for the in vitro diagnosis of Alzheimer's disease.

**Claims for the following Contracting States :**  
ES, GR

1. Use of the deoxyribonucleic acid according to Figures 1a - 1c and those functional equivalents of this sequence in which, due to the degeneration of the genetic code, individual nucleotides are exchanged or derivatised without this having an effect on the function of the nucleic acid, for the in vitro diagnosis of Alzheimer's disease. 5 10
2. Use of the deoxyribonucleic acid of the sequence from position 1 to position 2089 as shown in Figures 1a - 1c and those functional equivalents of this sequence in which, due to the degeneration of the genetic code, individual nucleotides are exchanged or derivatised without this having an effect on the function of the nucleic acid, for the in vitro diagnosis of Alzheimer's disease. 15 20
3. Use of the deoxyribonucleic acid of the sequence from position 600 to position 900 as shown in Figures 1a - 1c and those functional equivalents of this sequence in which, due to the degeneration of the genetic code, individual nucleotides are exchanged or derivatised without this having an effect on the function of the nucleic acid, for the in vitro diagnosis of Alzheimer's disease. 25 30
4. Use of the precursor protein of amyloid plaque protein of the sequence shown in Figures 1a - 1c and those functional equivalents of this sequence in which, due to the degeneration of the genetic code, individual amino acids are exchanged or derivatised without this having an effect on the function of the protein, for the in vitro diagnosis of Alzheimer's disease. 35 40
5. Use of the proteins or fragments according to Claims 1 to 4 as antigens for the production of antibodies. 45
6. Use of the antibodies according to Claim 5 for the in vitro diagnosis of Alzheimer's disease.

**Revendications**

Revendications pour les Etats contractants suivants : AT, BE, CH, DE, FR, GB, IT, LI, NL, SE

1. Acide désoxyribonucléique de séquence suivant les figures 1a-1c et équivalents fonctionnels de cette séquence dans lesquels, par dégénérescence du code génétique, certains nucléotides sont échangés ou dérivatisés sans que cela ait une influence sur la fonction de

l'acide nucléique.

2. Acide désoxyribonucléique de séquence allant de la position 1 à la position 2089 conformément aux figures 1a-1c et équivalents fonctionnels de cette séquence dans lesquels, du fait de la dégénérescence du code génétique, certains nucléotides sont échangés ou dérivatisés sans que cela exerce une influence sur la fonction de l'acide nucléique.
3. Acide désoxyribonucléique de séquence allant de la position 600 à la position 900 sur les figures 1a-1c et les équivalents fonctionnels de cette séquence dans lesquels, du fait de la dégénérescence du code génétique, certains nucléotides sont échangés ou dérivatisés sans que cela exerce une influence sur la fonction de l'acide nucléique.
4. Utilisation de l'acide désoxyribonucléique suivant les revendications 1 à 3 pour le diagnostic in vitro de la maladie d'Alzheimer.
5. Kit analytique contenant l'acide désoxyribonucléique suivant les revendications 1 à 3.
6. Précurseur protéique de la protéine de la plaque amyloïde de séquence suivant les figures 1a-1c et les équivalents fonctionnels de cette séquence dans lesquels, du fait de la dégénérescence du code génétique, des amino-acides individuels sont échangés ou sont dérivatisés sans que cela exerce une influence sur la fonction de la protéine.
7. Protéine de la séquence allant de la position 210 à la position 290 conformément aux figures 1a-1c et les équivalents fonctionnels de cette séquence dans lesquels, du fait de la dégénérescence du code génétique, des amino-acides individuels sont échangés ou dérivatisés sans que cela exerce une influence sur la fonction de la protéine.
8. Fragments des protéines suivant les revendications 6 et 7.
9. Utilisation des protéines ou des fragments suivant les revendications 6 à 8 pour le diagnostic in vitro de la maladie d'Alzheimer.
10. Utilisation des protéines ou des fragments suivant les revendications 6 à 8 comme antigènes non thérapeutiques pour la préparation d'anticorps.

11. Anticorps dirigés contre les protéines ou fragments suivant les revendications 6 à 8.

12. Kit analytique contenant les anticorps suivant la revendication 11.

5

13. Utilisation des anticorps suivant la revendication 11 pour le diagnostic in vitro de la maladie d'Alzheimer.

10

**Revendications pour les Etats contractants suivants : ES, GR**

1. Utilisation de l'acide désoxyribonucléique suivant les figures 1a-1c et équivalents fonctionnels de cette séquence dans lesquels, sous l'effet de la dégénérescence du code génétique, certains nucléotides sont échangés ou dérivatisés sans que cela ait une influence sur la fonction de l'acide nucléique, pour le diagnostic in vitro de la maladie d'Alzheimer.

15

20

2. Utilisation de l'acide désoxyribonucléique de la séquence allant de la position 1 à la position 2089 suivant les figures 1a-1c et des équivalents fonctionnels de cette séquence dans lesquels, du fait de la dégénérescence du code génétique, certains nucléotides sont échangés ou dérivatisés sans que cela exerce une influence sur la fonction de l'acide nucléique, pour le diagnostic in vitro de la maladie d'Alzheimer.

25

30

3. Utilisation de l'acide désoxyribonucléique de la séquence allant de la position 600 à la position 900 sur les figures 1a-1c et des équivalents fonctionnels de cette séquence dans lesquels, à cause de la dégénérescence du code génétique, certains nucléotides sont échangés ou dérivatisés sans que cela ait une influence sur la fonction de l'acide nucléique, pour le diagnostic in vitro de la maladie d'Alzheimer.

35

40

4. Utilisation du précurseur protéique de la protéine de la plaque amyloïde de la séquence suivant les revendications 1a-1c et des équivalents fonctionnels de cette séquence dans lesquels, à cause de la dégénérescence du code génétique, certains amino-acides sont échangés ou dérivatisés sans que cela exerce une influence sur la fonction de la protéine, pour le diagnostic in vitro de la maladie d'Alzheimer.

45

50

5. Utilisation des protéines ou des fragments suivant les revendications 1 à 4 comme antigènes pour la production d'anticorps.

55

6. Utilisation des anticorps suivant la revendication 5 pour le diagnostic in vitro de la maladie d'Alzheimer.

## EP 0 276 723 B1

AOTTTCTC00CA0C0GTAG0C0A0A -121  
 0CAC0CGGA0GAGCGT0C0C0G0GCCCC00G0A0C00C00C0GT0C0C0C0C00G0CA0A0 -61  
 CA0G0AC0C00C0GATCCCACTCGCAC0C0C0C0CACTC0GT0CCCC0C0CA00GT0C0C0 -1  
 AT0CT0CCCC00TTT00CACT0CTCCT0CT00CC0CCT00AC00CTC000C0CT00A00TA 60  
 M L P G L A L L L L A A H T A R A L E U  
 1 10 20  
 CCCACTGAT0GTAAT0CTG0CCT0CTG0CT0AACCCCA0ATT0CCAT0TTC0T0G0CA0A 120  
 P T D G N A G L L A E P Q I A H F C G R  
 30 40  
 CTGAACATGCACATGAATGTCCAGAAT0GGA0GT00GATTCA0ATCCATCA00GACCAAA 180  
 L N H H M N U Q N G K W D S D P S G T K  
 50 60  
 ACCTGCATT0ATACCAAGGAAGGCATCCTGCA0TATT0CCA0GAAGTCTACCCTGA0CT0 240  
 T C I D T K E G I L O Y C Q E U Y P E L  
 70 80  
 CAGATCACCAAT0TGGTAGAAGCCAA0CCCA0GTGACCATCCAGA0ACT0GTGCA0G0C00 300  
 Q I T N U U E A N Q P U T I Q N W C K R  
 90 100  
 GGCCGCAAGCAGT0CAAGACCCATCCCCACTTTGTGATTCCCTACCGCTGCTTAGTT0GT 360  
 G R K Q C K T H P H F U I P Y R C L U G  
 110 120  
 GAGTTTGTAA0GTATGCCCTTCTCGTTCCT0ACA0GT0CAA0ATTCTTACACCAGGAGA00 420  
 E F V S D A L L U P D K C K F L H Q E R  
 130 140  
 ATGGATGTTTGC0AA0CTCATCTTCACTGGC0AC0CGT0GCCAAAGAGACATGCAGTGA0 480  
 H D V C E T H L H W H T V A K E T C S E  
 150 160  
 AAGAGTACCA0CTTG0CATGACTACGG0CATGTTGCTGCCCTGC0GA0ATTGACAAGTTCCGA 540  
 K S T N L H D Y G H L L P C G I D K F R  
 170 180  
 0GGGTAGAGTTTGTGTGTTGCCCACTGGCTGAA0AAAGTGACAATGTGGATTCT0CT0AT 600  
 G U E F V C C P L A E E S D N U D S A D  
 190 200  
 0CGGA0CA00AT0ACTC0GATGTCT0GTGG00CGGAGCAGACAC0ACTAT0CAGAT000 660  
 A E E D D S D U W H G G A D T D Y A D G  
 210 220  
 AGT0AA0ACAAAGTAGTAGAAGTAGCAG00A0GAA0A0GT00CT0A00GT00AA0AA0AA 720  
 S E D K V U E U A E E E E U A E U E E E  
 230 240  
 0AAGCCGATGATGACGAGGACCATGAGGAT0GT0ATGAGGTAG00AA0A00GT0A0GAA 780  
 E A D D D E D D E D G D E U E E E A E E  
 250 260  
 CCCTACGAA0AGCCACAGAGACAACCACCAGCATT0CCACCACCACCACCACCACCACA 840  
 P Y E E A T E R T T S I A T T T T T T T  
 270 280  
 0AGTCTGTGGAAGAGGTGGTTCAGTTTCCTACAACAGCAGCCAGTACCCCTGATGCC0TT 900  
 E S U E E U U R U P T T A A S T P D A U  
 290 300  
 GACAAGTATCTC0AGACACCTGGGGATGAGAATGAACATGCCCATTTCCAGAAAGCCAAA 960  
 D K Y L E T P G D E N E H A H F Q K A K  
 310 320

FIG. 1a

EP 0 276 723 B1

GAGGGCTTGAGGCCAAGCACCAGAGAGAATGTCCCAGGTCATGACAGAATGGGAGAG 1020  
E R L E A K H R E R H S Q V H R E H E E  
330 340

GCAGAACGTCAGCAAGCAACTTGCCCTAAAGCTGATAAGCAAGGCAGTTATCCAGCATTTC 1080  
A E R Q A K N L P K A D K K A V I Q H F  
350 360

CAGGAGAAAGTGCAATCTTTGGAAACAGGAAGCAGCCAAACGAGACAGCACTGCTGGAG 1140  
Q E K V E S L E Q E A A N E R Q Q L V E  
370 380

ACACACATGGCCAGAGTGGAAGCCATGCTCAATGACCGCCGCCGCTGGCCCTGCAGAAC 1200  
T H M A R V E A M L N D R R R L A L E N  
390 400

TACATCACCGCTCTGCAGGCTGTTCCCTCCTCGGCCCTCGTCACGTGTTCAATATGCTAAA 1260  
Y I T A L Q A V P P R P R H V F N H L K  
410 420

AAGTATGTCCGCGCAGAACAGCAAGGACAGACAGCACACCCTAAAGCATTTCGAGCATGT 1320  
K Y U R A E Q K D R Q H T L K H F E H V  
430 440

CGCATGGTGGATCCCAAGAAAGCCGCTCAGATCCCGTCCCAGGTTATGACACACCTCCGT 1380  
R H V D P K K A A Q I R S Q V H T H L R  
450 460

GTGATTTATGAGCCCATGAATCAGTCTCTCTCCCTGCTCTACAACTGCGCTGCAGTGGCC 1440  
V I Y E R M N Q S L S L L Y N V P A V A  
470 480

GAGGAGATTCAAGATGAAGTTGATGAGCTGCTTCAGAAAGAGCAAACTATTTCAGATGAC 1500  
E E I Q D E U D E L L Q K E Q N Y S D D  
490 500

GTCTTGGCCAAACATGATTAATGAACCAAGGATCAATTACGGAAGGATGCTCTCATGCCA 1560  
V L A N H I S E P R I S Y D N D A L H P  
510 520

TCTTTGACCGAAACGAAACCAACCCGTGGAGCTCCTTCCCGTCAATGACAGTTTCAGCCTG 1620  
S L T E T K T T V E L L P V N G E F S L  
530 540

GACGATCTCCAACCCGTGACATTCTTTTGGGCTGACTCTGTGCCAGCCAAACAGCAAAAC 1680  
D D L Q P H H S F Q A D S V P A N T E N  
550 560

GAGTTGAGCCTGTTGATGCCCCGCTGCTGCCAGCCAGGACTGACCACTCGACCAAGT 1740  
E V E P U D A R P A A D R Q L T T R P Q  
570 580

FIG. 1b



EP 0 276 723 B1

TCTGGTTGACATATCAACCGGAGCAGATCTCTGAAGTGACATGATGCAATTC 1800  
 S O L T H I K T E E I S E U K H D A E F 600  
 390

CGACATGACTCAAGCATGAGGTTTCATCATCAAAATTTGTTCTTTGCAAGATGTT 1860  
 R H D S O Y E U H H Q K L U F F A E D U 620  
 610

GGTTCAAACAAAGGTGCAATCATTGCACTCATGGTGGGCGGTGTTGTCATAGCGACAGTG 1920  
 O S N K Q A I I B L H U Q Q U U I A T U 640  
 630

ATCOTCATCACCTTGTTGATGCTGAAGCAAAACAGTACACATCCATTCATCATGGTGTG 1980  
 I U I T L U M L K K K Q Y T S I H H G U 660  
 630

GTGGAGTTGACCGCGCTGTCAACCCAGAGGAGCGCCACCTGTCCAAGATGCAGCAGAAC 2040  
 U E U D A A U T P E E R H L S K H Q Q N 680  
 670

GGCTACGAAATCCAACTACAAGTTCTTTGAGCAGATGCAGAACTAGACCCCGCCACA 2100  
 Q Y E N P T Y K F F E Q M Q N \* 690

GCAGCCTCTGAAGTTGCACAQCAAAACCATTGCTTCACTACCCATCGGTGTCCATTTATA 2160  
 GAATAATGTGGGAAGAAACAAACCCGTTTATGATTTACTCATTATCGCCTTTTGACAAC 2220  
 TGTGCTGTAACACAAGTAGATGCTTGAACCTGAAATTAATCCACACATCAGTAATGTATTC 2280  
 TATCTCTCTTTACATTTTGTCTCTATACTACATTATTAATGGGTTTTGTGTACTGTAA 2340  
 GAATTTAGCTGTATCAAACTAGTGCATGAATAGATTCTCTCCTGATTATTTATCACATAG 2400  
 CCCCTTAGCCAGTTGTATATTATTCTTGTGGTTTGTGACCCAATTAAGTCTACTTTACA 2460  
 TATGCTTTAAGAAATCGATGGGGGATGCTTCATGTGAACGTGGGAATTCAGCTGCTTCTCT 2520  
 TGCCTAAGTATTCCTTTCTGATCACTATGCATTTTAAAGTTAAACATTTTAAGTATTT 2580  
 CAGATGCTTTAGAGAGATTTTTTTTCCATGACTGCATTTTACTGTACAGATTGCTGCTTC 2640  
 TGCTATATTTGTGATATAGGAATTAAGAGGATACACACGTTTGTCTTCTCGTGCCTGTTT 2700  
 TATGTGCACACATTAGGCATTGAGACTTCAAGCTTTTCTTTTTTGTCCACGTATCTTTG 2760  
 GGTCTTTGATAAAAGAAAGAAATCCCTGTTCAATTGTAAGCACTTTTACGGGGCGGGTGGG 2820  
 AGGGGTGCTCTGCTGGTCTTCAATTACCAAGAAATTCTCCAAACAAATTTTCTGCAGGATG 2880  
 ATTGTACAGAAATCATTGCTTATGACATGATCGCTTCTACACTGTATTACATAAAATAAT 2940  
 TAAATAAAATAACCCCGGCAAGACTTTTCTTTGAAGGATGACTACAGCATTAAATAAT 3000  
 CGAAGTAATTTTGGTGGTGGGAGAGGAGGAGATTCAATTTTCTTTAACCAGTCTGAAGTT 3060  
 TCATTTATGATACAAAAGAGATGAAATGGAAATGGCAATATAAGGGGATGAGCAAGGC 3120  
 ATGCCTGACAAACCTTCTTTTAAGATGTGCTTCAATTTGTATAAAATGGTGTGTTTCA 3180  
 TGTAAATAATACATTCTTGGAGGAGC - poly(A) tail

FIG. 1c